5

MPEG Coding Principles

In this chapter we review some of the basic principles used in MPEG coding. We use this chapter to introduce some of the terminology used throughout the book, but our treatment here is necessarily brief. Readers who have no familiarity with data compression will probably want to supplement the treatment here with texts such as [RJ91], [BK95], [Say96], [Hof97] and [PM93].

We first present a very high-level view of a coding system, showing how the system can be divided into modeling and entropy coding sections. Then, we discuss the very important topics of entropy coding and coding models, in that order. Finally, we discuss the specifics of the MPEG-1 coding techniques and present block diagrams for MPEG-1 encoders and decoders. Note that our use of the term system in this chapter is unrelated to the MPEG system layer.

5.1 Coding system structure

The high-level coding system diagram sketched in Figure 5.1 illustrates the structure of a typical encoder and decoder system. The analog to digital conversion (A/D) determines the basic resolution and precision of the input data, and thus is a very important step in reducing the almost unlimited data that potentially is available from the original scene to a manageable level. However, data reduction does not necessarily stop once the digitization is completed.

Compression systems that do no further data reduction once the picture is digitized are lossless systems; these lossless compression systems rarely compress natural image data by more than a factor of 2 to 3. Compression systems such as MPEG need to achieve considerably more than an order of magnitude higher compression than this, and they do this by means of
further lossy data reduction after the digitization.

As shown in Figure 5.1, it is convenient to separate a coding system into two parts. The first part is the encoder model that performs lossy data reduction in the process of changing the digital source data into a more abstract representation which is conventionally labeled symbols. The second part then codes these symbols in a process that minimizes the bitstream length in a statistical sense. This second step is called entropy coding.  

5.1.1 Isolating the model

The decoder in Figure 5.1 reverses the encoding process, at least to the extent possible. It first losslessly converts the compressed data back to symbols, and then rebuilds a digital picture that is (hopefully) a visually close approximation to the original digital source data. This digital data is fed through a D/A (digital to analog) converter to recreate an analog output signal for the display.

For the moment consider the entropy encoder and decoder to be “black boxes” with the following properties: The entropy encoder accepts a stream of symbols from the encoder model and converts it to compressed data; the entropy decoder decodes that compressed data and returns an identical

---

1 In [PM93] a slightly different decomposition is used in which an intermediate stage of descriptors is defined. This was necessary because of the two entropy-coding techniques used in JPEG. MPEG uses only one type of entropy coding and this intermediate stage is not needed. We therefore use the more conventional decomposition in this book.